Progress in Biophysics and Molecular Biology 201 (2026) 341–349

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

# Exome-matched protein as the ideal dietary amino acid pattern: A narrative review

Rainer J. Klement [a,b,*]

[a] Department of Radiotherapy and Radiation Oncology, Leopoldina Hospital Schweinfurt, Schweinfurt, Germany
[b] SaluGena | Praxis für evolutionäre Naturheilkunde, Wipfeld, Germany

ARTICLE INFO

*Keywords:*
Amino acids
Evolution
Exons
LUCA
Genetics
Protein quality

ABSTRACT

With few exceptions, all organisms on Earth use a common amino acid alphabet for synthesizing protein that consists of 20 canonical amino acids. Heterotrophic organisms have to obtain these amino acids from their food, especially the indispensable amino acids. The amount of and ratio between the indispensable amino acids is therefore a major determinant of protein quality. There exist several measures of protein quality for human nutrition and all of them rely on the empirical estimation of amino acid requirements. With the availability of whole genome sequencing data, however, a new method of deductively deriving amino acid requirements became available which uses the information encoded on the exome (the totality of protein-coding exons). By translating an organism's exome into the corresponding proteome *in silico*, the average encoded amino acid pattern can be computed and used *ex hypothesi* for defining an ideal amino acid pattern. Here I review the theoretical concepts behind exome-matched proteins and summarize the preclinical data that provide preliminary evidence for the hypothesis that such proteins indeed constitute an ideal protein source to maximize growth, reproduction and health. While clinical trials are yet to confirm this hypothesis in humans, there is a broad hypothetical range of applications of exome-matched proteins for human consumption. This review concludes that the concept of exome-matched proteins is interesting theoretically, promising for practical applications in animal and human nutrition and stimulating for further transdisciplinary research.

## 1. Introduction

The quality of dietary protein plays a significant role for both human and animal nutrition and health. Protein quality is determined by three factors (Adhikari et al., 2022): Firstly, the digestibility, i.e., which fraction of the protein's amino acids (AAs) are actually absorbed within the small intestine. For example, on average only 70-90% of the AAs from plant proteins can be digested due to the antinutrients contained within plants; in the case of sweet potatoes this fraction is only 52% (Table 2.7 in Wu, 2022). Digestibility of animal proteins lies in the range between 95% and 100%. Secondly, protein quality depends on the overall AA profile or pattern, respectively. Ideally, the AA pattern would match the actual AA requirement of the organism without individual AAs lacking or being in excess. Thirdly, protein quality depends on the concentration of indispensable (also called essential) AAs, which are AAs that an organism cannot synthesize *de novo* and has to obtain through the diet. If only one indispensable AA is missing for the synthesis of a protein, this protein cannot be produced, even if all other AAs are in excess. In the end, because the body has very limited ability to store them, excessive AAs undergo catabolic deamination reactions and are oxidized,[1] which leads to ammonium production, its detoxification in the urea cycle and subsequent nitrogen excretion in the urine (Wu, 2022).

The currently favored measure of protein quality for humans is the digestible indispensable amino acid score (DIAAS). For an individual indispensable AA $x$, its $DIAAS_x$ (in %) is calculated as (FAO, 2013)

$$DIAAS_x\ [\%] = 100 \times \frac{\text{mg of digestible AA } x \text{ in 1 g of the dietary protein}}{\text{mg of the same AA in 1 g of the reference protein}} \quad (1)$$

The digestible amount of an AA $x$ is thereby given by the total amount of $x$ (in mg per g crude protein) × the standardized ileal digestibility (Herreman et al., 2020). The DIAAS of a protein source is then given by

* Klinik für Strahlentherapie, Leopoldina Krankenhaus Schweinfurt, Robert-Koch-Straße 10, 97422, Schweinfurt, Germany.
*E-mail address:* rklement@leopoldina.de.

[1] This is the basic principle behind the indicator amino acid oxidation method to determine the requirement of essential AAs (Elango et al., 2008).

https://doi.org/10.1016/j.pbiomolbio.2026.08.001
Received 22 January 2026; Received in revised form 13 May 2026; Accepted 2 August 2026
Available online 3 August 2026
0079-6107/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

$$\mathrm{DIAAS}_{\mathrm{Protein}} = \min_{\forall x} \mathrm{DIAAS}_x \quad (2)$$

The indispensable AA with the lowest DIAAS is called the limiting indispensable AA. It can be seen from Eq. (1) that, *ceteris paribus*, the DIAAS depends on the selection of the reference protein that is supposed to reflect the ideal AA pattern and thereby determines the individual indispensable AA requirements. The FAO has defined three different reference protein patterns: one for infants (<6 months of age) based on the AA pattern of human breast milk, one for children (0.5-3 years) based on the reference AA pattern of 6 month old infants, and one for older children, adolescents and adults based on the reference AA pattern of 3-10 year old children (FAO, 2013). The reference AA patterns for 6 month old infants and 3-10 year old children in the FAO report were taken from a 2007 WHO/FAO/UNU report in which they were *inductively* derived from an amalgamation of empirical AA requirement estimations using both nitrogen balance and isotope tracer methods (World Health Organization, 2007). Table 1 exemplarily shows the DIAAS for several foods compiled from the study of Herreman et al. (2020) based on the reference pattern of 3-10 year old children.

However, empirical methods to derive AA reference patterns have their shortcomings and provide only approximations to the real AA requirement. In addition, it is known that dispensable (also called non-essential) AAs are also important individual micronutrients in food proteins (Hou et al., 2015; Hou and Wu, 2017), but their individual requirements are less well defined by empirical methods. It would be of great importance if a method exists that allows to *deductively* deriving optimal AA requirements based on first principles of physiology. Piper et al. have laid the groundworks of such a method by showing that an AA pattern matched to the average AA composition of the translated protein-coding exome of *Drosophila melanogaster* was superior to traditional food sources with respect to several metrics of protein quality assessment (Piper et al., 2017). These metrics included fecundity, growth, nitrogen balance, satiety and longevity (MacArthur and Mitchell, 2017). Meanwhile, the concept of exome-matched proteins (EMPs) as the ideal species-specific AA pattern has been developed and tested further (Gong et al., 2024; Ma et al., 2022; Ortega et al., 2023; Wu et al., 2024) and an EMP composition for humans is available (Klement, 2024). In this review, I will provide a transdisciplinary narrative of the hypothesis that an EMP ideally covers an organism's AA requirements. The review covers topics such as genome evolution, what constitutes an organism's exome, the principles of an EMP and the preclinical evidence supporting the above hypothesis. It thereby utilizes three conceptually distinct types of inference corresponding to distinct levels of organization that are studied by different scientific disciplines:

- Descriptive inferences from the molecular level of AAs, the genome, transcriptome and proteome; this involves evolutionary and molecular biology, biochemistry, genetics and computer science.
- Modeling inferences based on the level of organs and whole organisms with which preclinical studies are concerned; this involves disciplines such as molecular biology, pathology, physiology and nutritional sciences.
- Normative inferences, i.e., how we should translate the data collected on the previously described levels to the higher levels of organization involving human individuals, particular social groups, pets and livestock and therefore particular parts of the economy, for example by potentially proposing new dietary recommendations and supplements with the aim of improving feed efficacy and health; normative inferences involve disciplines such as nutritional sciences, economics and veterinary and human medicine.

**Table 1**
Digestibility-corrected indispensable amino acid score for several protein sources with the reference pattern of 3-10 year old children.

| Protein | DIAAS (limiting amino acid) |
|---|---|
| Casein | **137** ± 5.9 (NA) |
| Pork | **126** ± 9.7 (NA) |
| Chicken egg | **111** ± 17.5 (NA) |
| Potato | **125** ± 9.1 (NA) |
| Whey protein | **106** ± 13.5 (NA) |
| Soy protein | **103** ± 7.9 (NA) |
| Pea protein | **83** ± 14.4 (Met + Cys) |
| Oat | **68** ± 6.8 (Lys) |
| Wheat | **56** ± 13.8 (Lys) |
| Gelatin | **4** ± 7.7 (Trp) |
| Pea + Wheat mixture | Max. **93** (for a 40/60 mixture) |

## 2. Genome evolution and the amino acid alphabet

Our Solar system emerged approximately 4.567 Ga ago in the course of a star-forming event possibly triggered by the transition of a small satellite galaxy, the Sagittarius dwarf galaxy, through the Milky Way's disk (Ruiz-Lara et al., 2020). This time estimate is derived from the chemical composition of meteorites called chondrites, which are remnants from the protoplanetary disk that gave rise to the inner and outer planets of the Solar system (Halliday and Canup, 2023). Carbonaceous chondrites in particular not only have largely preserved the composition of the protoplanetary disk, but also contain many hints for the presence of AAs (e.g. alanine, glycine and valine) and other organic molecules as possible precursors, or even signatures, of early life forms (Glavin et al., 2006; Shimoyma and Ogasawara, 2002). This, together with other evidence for AA abundance in star-forming regions,[2] suggests that AAs would have principally been available on the primordial Earth as the building blocks of the first cellular organisms. It is possible that these first primordial organisms used an early AA alphabet consisting of a small set of abiotic AAs, and that other "late" AAs were subsequently added to the alphabet after becoming available through biosynthetic production.[3] However, the reason for and the chronological sequence of selection of AAs into the AA alphabet and their encoding on RNA/DNA molecules are still debated (Makarov et al., 2023; Wehbi et al., 2024). What is also uncertain is the time point at which life using a genetic code emerged on Earth – according to some authors, it should have been present from the very beginning (Steele et al., 2019; Wickramasinghe, 2018), while others argue for an abiogenic metabolism-first scenario (Martin, 2020).

In any case, more accurate information can be obtained about the last universal common ancestor (LUCA) that can be studied with standard phylogenetic methods (Goldman and Becerra, 2024). Moody et al. (2024) have estimated that the LUCA lived ~ 4.2 Ga (4.09-4.33Ga) ago, which is only a few hundred million years after the formation of the proto-Earth. By definition, the LUCA should have had a genetic code very similar to that which is shared (with some exceptions) by all extant life forms currently inhabiting Earth. This code consists of

[2] Such evidence stems from observations including (i) the indirect detection of glycine in the coma of comets (Altwegg et al., 2016; Elsila et al., 2009); (ii) laboratory experiments confirming the possibility of *de novo* production of AAs and peptides within molecular clouds *in glacies* both independent (Ioppolo et al., 2021; Krasnokutski et al., 2022; Shivani et al., 2017) and dependent on UV-radiation (Munñoz Caro et al., 2002; Woon, 2002); (iii) evidence for the presence of phenylalanine and tyrosine in the star-forming region IC348 (Iglesias-Groth et al., 2025). In addition, a wealth of amino acids on the early Earth could have been produced *in situ* through energetic processes such as atmospheric spark discharges or in hydrothermal vents (reviewed by Kirschning, 2022).

[3] See, however, Martin (2020) for an alternative view according to which an enzyme-independent and acellular primitive metabolism developed first at hydrothermal vents, with AA and peptide production as a byproduct.

triple-sequences (codons) of the four nucleotides adenosine (A), thymine (T), guanine (G) and cytosine (C), whereby each of the $4^3 = 64$ possible codons encodes one of the 20 canonical proteinogenic AAs[4] (the three codons UAA, UAG and UGA are an exception, as they are used as stop signals during protein synthesis in most organisms) (Koonin and Novozhilov, 2017).

Moody et al. (2024) estimated that the LUCA had a genome of 2.75 Mb (2.49-2.99 Mb), encoding around 2600 proteins. Overall, Moody et al. depict the LUCA as a unicellular organism with a DNA genome of similar complexity as that of modern-day prokaryotes (Goldman and Becerra, 2024).

It is likely that the LUCA genome already contained mobile introns[5] with self-splicing RNA activity (rybozomes) from which the later spliceosomal introns in eukaryotes emerged (Poverennaya and Roytberg, 2020; Rogozin et al., 2012). These ancient introns still persist as group II introns in the genome of recent prokaryotes and the organelles of eukaryotes. Eukaryotes likely emerged concurrently with or after the great oxygenation event some 2.5-2.1 Ga ago by a symbiogenetic fusion process between an Asgard archaeon and alpha-proteobacteria (Appler et al., 2026). The latter are considered the ancestors of mitochondria. Eugene Koonin (2009) estimated that the last eukaryotic common ancestor (LECA) possessed up to 70-80% introns in its nuclear genome stemming from invading group II introns of the bacterial symbiont. Zimmerly and Semper (2015) summarized the hypothesis of spliceosomal intron evolution as follows:

> *At some point prior to the LECA, group II introns likely invaded the nuclear genome and proliferated as mobile DNAs. The invading group II intron(s) could have come from the genome of the alpha-proteobacterium that became the mitochondrial endosymbiont or alternatively could have been transferred from a bacterium to the nuclear genome after establishment of the mitochondrion.*

Many studies have investigated the properties and evolution of spliceosomal introns and exons in eukaryotes since the LECA. A main insight is that since the LECA, both intron losses and gains were substantial, but losses generally outnumbered gains, the latter often occurring in connection with major evolutionary transitions towards more complex organisms such as the emergence of the Metazoa (Rogozin et al., 2012). Indeed, intron density is generally higher and exon density accordingly lower in animals and plants compared to fungi and unicellular eukaryotes. Duan and Cao (2025) estimated that the human genome has an exon density of $\sim$ 4.4% compared to $\sim$ 5.7% in mice and $\sim$ 27% in the fruit fly (*Drosophila melanogaster*), thereby considering all exons including those in untranslated regions and non-coding RNAs. However, exon density in humans is $\sim$ 1% if only internal protein-coding exons are considered (Stepankiw et al., 2023). Across all 23 human chromosomes, there are on average 8-10 exons and 7-9 introns per gene[6] (Sakharkar et al., 2004); the median exon length was estimated as $\sim$ 140 bp which is shorter than the median intron length by a factor of $\sim$ 10 (Duan and Cao, 2025).

While it is important to consider that not all exons are located within protein-coding regions (Aspden et al., 2023), we are here mainly concerned with those that are. In the following, we will therefore use the words exon and exome (the totality of exons) under the assumption that they refer to the protein-coding regions. With the availability of whole genome sequencing data, the possibility has emerged to translate the exome of a species into a predicted proteome *in silico*. Such a set of hypothetically synthesized proteins allows the computation of an average AA pattern, which is species-specific and may, *prima facie* and on a descriptive level, correspond to an ideal dietary AA pattern for that species (Table 2). In recent years, several preclinical studies have tested this hypothesis, with promising findings. These are summarized in the following section and represent the modelling level of inference.

## 3. Preclinical studies confirming the superiority of exome-matched diets over tradition protein sources

### 3.1. An exome-matched protein diet optimizes growth, fecundity, satiety and nitrogen balance in fruit flies and mice

Whether an EMP constitutes an ideal protein source has first been tested in a seminal study by Matthew Piper and colleagues that used fruit flies and mice as model organisms (Piper et al., 2017). The EMP was derived as the average relative proportion of each AA based on its occurrence across the entire *in silico* translated exome; this pattern differed significantly from the AA composition found in natural or standard laboratory proteins. Using chemically defined (holidic) diets, Piper et al. demonstrated that, compared to two non-matched protein diets, the EMP diet reduced uric acid production,[7] increased satiety,

**Table 2**
Exome-derived amino acid patterns for humans and mice.

| Amino acid | Exome-derived amino acid pattern for humans | Exome-derived amino acid pattern for mice |
|---|---|---|
| Leucine (L, Leu) | 9.732% | 9.735% |
| Serine (S, Ser) | 8.689% | 8.808% |
| Glutaminc acid (E, Glu) | 7.345% | 7.150% |
| Alanine (A, Ala) | 6.800% | 6.836% |
| Proline (P, Pro) | 6.426% | 6.459% |
| Glycine (G, Gly) | 6.316% | 6.323% |
| Lysine (K, Lys) | 5.981% | 5.815% |
| Valine (V, Val) | 5.813% | 5.881% |
| Arginine (R, Arg) | 5.644% | 5.750% |
| Threonine (T, Thr) | 5.347% | 5.406% |
| Glutamine (Q, Gln) | 4.987% | 4.954% |
| Aspartic Acid (D, Asp) | 4.835% | 4.836% |
| Isoleucine (I, Ile) | 4.281% | 4.187% |
| Asparagine (N, Asp) | 3.697% | 3.576% |
| Phenylalanine (F, Phe) | 3.461% | 3.496% |
| Histidine (H, His) | 2.669% | 2.684% |
| Tyrosine (Y, Tyr) | 2.549% | 2.591% |
| Methionine (M, Met) | 2.156% | 2.187% |
| Cysteine (C, Cys) | 2.135% | 2.187% |
| Tryptophan (W, Trp) | 1.135% | 1.138% |

Data are taken from Gong et al. (2024) who translated the human and mouse exome from the NCBI genome database. Briefly, the percentages represent the average composition of 123,186 protein sequences from the human genome and 87,205 protein sequences from the mouse genome (sequences shorter than 100 or greater than 2000 amino acids were excluded from the computation).

[4] Why exactly these 20 canonical amino acids and not others were encoded by the genome of the LUCA is an interesting question. One answer given to this question is that this amino acid alphabet has been optimized for building a broad variety of proteins, because the 20 canonical amino acids occupy the possible ranges of the three fundamental amino acid properties size, charge and hydrophobicity more evenly and broadly than other possible amino acid combinations (Mayer-Bacon and Freeland, 2021; Philip and Freeland, 2011). Another study found that inclusion of other abiotic amino acids than the canonical ones into a proteinogenic amino acid alphabet would have resulted in polypeptides that are oversolubilized and have low folding efficiency (Makarov et al., 2023).

[5] Introns are defined as DNA sections that are removed during RNA splicing, while sections that are retained and result in a mature mRNA transcript are termed exons.

[6] Because introns are the segments between exons, their number is equal to the number of exons minus 1.

[7] Adult flies excrete nitrogenous waste by converting ammonia into uric acid.

promoted growth, and enhanced reproduction measured by female egg laying capacity in *D. melanogaster*. Exome matching allowed the authors to identify the limiting indispensable AA in each non-matched protein and to show that egg production was reduced proportional to the relative lack of this limiting AA in the mismatched compared to the EMP (Piper et al., 2017). Intriguingly, Piper et al. found that flies fed EMP diets only differing in the amount of the holidic protein displayed a single and relatively low protein intake optimum at 10.7 g/L for simultaneously maximizing growth, egg laying capacity and lifespan. In contrast, mismatched diets could either optimize lifespan at low protein concentrations (10.7 g/L) or growth and reproduction at high protein concentrations (≥21.4 g/L). Thus, it was possible to circumvent the well-known longevity-reproduction tradeoff observed with conventional protein sources by using an EMP source at low concentration. Finally, the results were qualitatively confirmed in mice: compared to mice fed casein or another mismatched protein source, mice fed with EMP exhibited significantly reduced urinary nitrogen excretion alongside significantly higher lean and fat mass gains and increased bone thickness and mineral density, respectively. In summary, this study showed that matching protein sources to the species-specific translated exome has the potential for optimizing nitrogen utilization and health in both flies and mice.

### 3.2. Exome matching reduces the amount of protein needed for optimal reproduction

In a 2022 follow-up study, the same research group used an *D. melanogaster* egg-laying assay to examine the differential effects of protein quality and quantity on female reproduction (Ma et al., 2022). The researchers created 40 different holidic diets of which half used an EMP as the sole protein source and the other half a mismatched protein source. For each of these two protein sources, diets were generated that spanned a range of five different protein:carbohydrate ratios and – for a fixed macronutrient ratio – four different caloric densities, resulting in 20 diets with EMP and 20 diets with a mismatched protein source that varied in the total amount of protein and carbohydrates. The experiments showed that for a fixed amount of protein and carbohydrate, female flies feeding on EMP always laid more eggs than flies feeding on the mismatched protein. In both cases, reproductive output was not significantly dependent on carbohydrate intake. However, while flies feeding on mismatched protein diets increased their reproductive output linearly with the amount of protein consumed, flies feeding on EMP diets maximized their reproductive output on intermediate amounts of protein at ∼ 15 g/L, which was ∼ 2.5 times less protein than required for flies on a mismatched diet to achieve the same output. Thus, this study pointed out the importance of protein quality alongside quantity for determining health-related outcomes in fruit flies.

### 3.3. Adjusting for transcription levels of genes does not provide further benefits to exome-matched proteins in flies

One year later, Ortega et al. (2023) showed that EMP was equally effective as an exome-based AA pattern further tailored to the sex-specific gene expression differences between male and female flies. To this aim, two additional holidic proteins were created besides EMP: one protein in which the individual AA proportions reflected the relative proportion of each AA in an average male transcriptome (denoted as MALEAA), and one in which the individual AA proportions reflected the relative proportion of each AA in an average female transcriptome (FEMALEAA). Using a similar egg laying assay as used before by Piper et al. (2017), the research group showed that female reproductive output was similar between EMP and FEMALEAA diets and only determined by the absolute amount of protein. By comparing the AA pattern between EMP and FEMALEAA, lysine (K, Lys) was identified as the potential limiting indispensable AA in EMP. By feeding female flies with 10 g/L EMP in which only lysine was diluted to a certain percentage of its original concentration, Ortega et al. showed that egg production was reduced proportional to the dilution level of lysine. Thus, it was concluded that flies were not able to compensate for lysine deficiency by mobilizing it from body stores, so that the finding that EMP and FEMALEAA produced a similar reproductive output was not due to such a compensatory mechanism. In an analogous experiment with male flies in which the ability to fertilize virgin females was tested, it was shown that for a fixed amount of protein, EMP and MALEAA resulted in indistinguishable results. Here, tryptophan (W, Trp) was identified as the limiting indispensable AA in EMP (under the assumption that MALEAA is the ideal AA pattern), but again the fact that EMP performed equally well as MALEAA was not due to Trp mobilization from body tissues. A possible explanation of these findings is that the overall AA pattern of EMP, MALEAA and FEMALEAAs looks very similar and that the small deviations in the relative proportion of each AA were too small to lead to measurable differences in reproduction. A similar AA pattern was also found between EMP and the transcriptome-matched AA pattern of flies challenged with *E. coli* infection (Ortega et al., 2023). Thus, the key finding of this work was that although biological sex theoretically predicts varying requirements for individual AAs, the primary difference in practice lies merely in the absolute quantity of protein required, rather than the adjustment for sex-specific gene expression — the ideal AA pattern appears to correspond fundamentally to the species-specific EMP. It would be intriguing to test the logically following hypothesis that the EMP remains the ideal AA pattern regardless of age, physical activity or other factors and that in these cases, too, only the absolute protein requirement changes (as protein needs increase both with age and heightened physical activity, see Bauer et al., 2013; Burd et al., 2019).

### 3.4. Exome-matched proteins can be designed from natural food proteins

Finally, the same research group demonstrated in a fourth study that mice under protein restriction showed significantly better growth and more efficient anabolic utilization of dietary AAs when fed a diet matching the mouse-specific EMP (Wu et al., 2024). Notably, this study did not rely solely on a mixture of synthetic AAs; instead, the EMP was designed primarily through a computer-optimized combination of three different milk protein powders, supplemented with only minimal amounts of individual AAs to further align with the exome-matched AA pattern. The optimization algorithm[8] also allowed the design of an EMP diet that included an additional adjustment for digestibility of the individual protein ingredients (using data from in vitro digestibility simulations by Wu et al., 2017). The additional digestibility adjustment increased body mass gains and feed conversion efficiency further than the unadjusted EMP diet. Finally, the testicular daily sperm production rate was higher in male mice fed with either the digestibility-adjusted or unadjusted protein blend, although other measures of reproductive function were not significantly different between male and female mice fed either EMP or mismatched protein diets (Wu et al., 2024). In summary, this study provided a proof-of-principle that EMP diets designed from natural food ingredients have the potential to increase feed conversion efficiency in animals, which could be especially relevant for the veterinary industry.

### 3.5. An exome-matched diet supports tumor-bearing mice

Besides Matthew Piper's research group, one independent research group from China conducted a preclinical study in mice to test the potential benefits of an EMP diet. Gong et al. (2024) used the murine 4T1 breast cancer model and divided tumor-bearing mice into four different

[8] This algorithm named ePROB (“exome-informed PROtein Balancing”) is now under consideration to re-design foods for human consumers or the veterinary industry (Bennett et al., 2023).

groups, all of which received paclitaxel chemotherapy but differed in their protein supplementation. One group received a nutritional supplement developed for cancer patients at a dose of 1.8 g/kg/day. A second group received the same preparation at 1.8 g/kg/day, but adjusted by adding the underrepresented AAs serine and glycine to match the murine EMP (Table 2) as closely as possible. A third group received the unadjusted protein at double the dose (3.6 g/kg/day). The fourth group received no protein supplementation in addition to their standard diet. I have already summarized the results of this study in an Editorial for the journal *Translational Breast Cancer Research* (Klement, 2024).

First, it is important to note that supplemental protein — regardless of composition or dose — did not promote tumor growth compared to the group that received no protein preparation. However, the mice receiving supplemental protein exhibited higher body weight during the third and final week of the experiment. This is reassuring, given that clinical guidelines recommend higher protein intakes in the range of 1.5–2 g/kg/day for cancer patients to compensate for systemic inflammation and tumor-induced muscle protein degradation, without fearing that additional AAs might "feed" the tumor (Arends et al., 2017). In fact, cancer cells are largely independent of dietary AA intake, as they frequently utilize glutamine as an energy source (the plasma levels of which are barely influenced by diet; see Seyfried and Chinopoulos, 2021), or they cannibalize their host through a process called macropinocytosis, in which whole peptides and proteins are engulfed by the cell and subsequently broken down into AAs to provide anabolic and catabolic (energy-producing) substrates (Commisso et al., 2013).

A second key observation was that mice receiving the exome-adjusted protein supplement were able to increase their grip strength (an indicator of muscle strength) significantly more than mice receiving either no supplement or the unadjusted, glycine- and serine-poor protein preparation. However, the group receiving the latter at double the dose was able to increase its grip strength to a similar extent as the mice receiving the EMP-adjusted formula. Subsequent transcriptomic profiling of the skeletal muscle tissue showed that the expression of the Complement 3 (C3) protein was significantly upregulated in the muscle cells of mice receiving the EMP-adjusted protein. C3 is the central activator of the complement system and has been proven to promote muscle regeneration (Zhang et al., 2017).

A third interesting result of the study was that protein supplementation counteracted the intestinal dysbiosis caused by paclitaxel. Interestingly, the strongest effect here was observed with the double dose of the unadjusted protein supplement, suggesting that the quantity of AAs might be more critical for restoring chemotherapy-induced dysbiosis than protein quality. However, there appeared to be additional benefits of aligning AAs with the EMP, as the adjusted protein supplementation group showed a significantly higher relative abundance of bacteria of the genus Alistipes compared to the unadjusted group; Alistipes bacteria are considered beneficial in this context. This finding is interesting but is currently insufficient to draw clear conclusions regarding the impact of EMP on the microbiome in mice.

## 4. Translational implications

Taken together, the preclinical studies summarized above provide preliminary evidence that the ideal AA pattern for a given species is anchored in the genome. The concept of exome matching allows a deductive inference of this ideal AA pattern, in contrast to inductive inferences based on empirical methods such as nitrogen utilization or tracer-based studies. By discussing the potential translational implications of these findings, we can now discuss certain normative inferences. I will thereby focus on three different implications:

First, EMPs could serve as a new reference for evaluating the quality of natural protein sources. The adjustment for typical age-specific gene expression profiles, analogous to the sex-specific adjustments performed by Ortega et al. (2023), could allow further refinements and an age-dependent specification of protein quality scores, similar to the DIAAS (FAO, 2013). In Figs. 1 and 2 I have plotted the ratio between the concentration of individual AAs in common animal and plant foods, respectively, and their concentration in the human EMP.[9] Fig. 2 shows that the AA with the lowest concentration relative to EMP is an indispensable AA for all displayed plant proteins, while it is a dispensable AA in animal proteins, as shown in Fig. 1. As a comparison between Figs. 1 and 2 shows, the human EMP is significantly more similar to animal proteins than to plant proteins, particularly regarding the proportion of indispensable AAs (shown as red triangles), of which some are found in far too low concentrations in plants (<70% of EMP). The whole chicken egg, for instance, comes relatively close to the EMP (Esumi, 2025), which aligns with the biological value of 100 — a metric derived purely empirically and established somewhat arbitrarily in the early 20th century using the egg as a reference protein. Several studies exist regarding the high protein quality of eggs. These include a randomized controlled trial in infants aged 6–9 months, which showed accelerated growth through the daily supplementation of one egg (Iannotti et al., 2017), and a case series of eight patients with severe burns who achieved rapid normalization of their serum protein levels by consuming 35 eggs daily (Hirshowitz et al., 1975). Due to its relative proximity (relative in terms of the AA composition distance; see Fig. 1 and Esumi, 2025) to the chicken egg and other animal protein sources, the concept of an EMP as an ideal AA pattern is thus indirectly confirmed through empirical observations.

The second important aspect, if EMP formulations would indeed represent ideal protein sources, is the potential for the animal feed industry and veterinary medicine. It is long known in livestock management that improving the match between protein or AAs fed to animals and the actual AA requirements reduces nitrogen waste (Rotz, 2004). Feeding animals according to their species-specific exome-matched AA pattern could potentially minimize nitrogen excretion while simultaneously maximizing animal growth and health.

A third translational implication is the potential for the dietary supplement market and human health, a point also noted in a 2017 editorial on the study by Piper et al. (MacArthur and Mitchell, 2017). The idea of an ideal AA pattern has already been the basis for a dietary supplement known as the "Master Amino Acid Pattern" (MAP), an AA formula consisting of all human indispensable AAs except for histidine. Since its patent ran out, MAP is sold under various names, but all contain the same proportion of indispensable AAs that – according to the claims – minimize nitrogen waste due to a high net nitrogen utilization of 99%[10] This number was derived in a double-blind triple-crossover study by Maurizio Lucà-Moretti who fed 66 subjects (33 males, 33 females) with one of three protein sources (MAP, whole egg protein or the AA formula of whole egg protein) besides other foods containing only negligible amounts of protein. After an initial 30-day acclimatization diet "to equalize and stabilize subjects' protein and energy metabolism, thus avoiding metabolism differences that could affect their N balance "(Lucà-Moretti, 1998), each diet was given consecutively in a random order over 28 days, delivering 0.4g protein per kg per day and 50 kcal/kg/day. The study found that only MAP resulted in an equilibrated nitrogen balance of 0.003 ± 0.001, while the other two diets resulted in a negative nitrogen balance. The net nitrogen utilization which corresponds to the utilization of AAs for anabolic rather than catabolic (deamination) pathways was estimated as 99% for MAP and 68% for whole hen protein. However, there are some critical points to mention regarding the concept of MAP as the "ideal amino acid formula for human nutrition" (Lucà-Moretti, 1998). Firstly, it must be pointed

[9] The calculations of these ratios are analogous to Eq. (1), except that the ratios are also computed for the dispensable AAs and digestibility is not considered.

[10] See for example https://www.gesundheitsmanufaktur.de/map-aminosaeuren-master-amino-acid-pattern, accessed on May 12th, 2026.

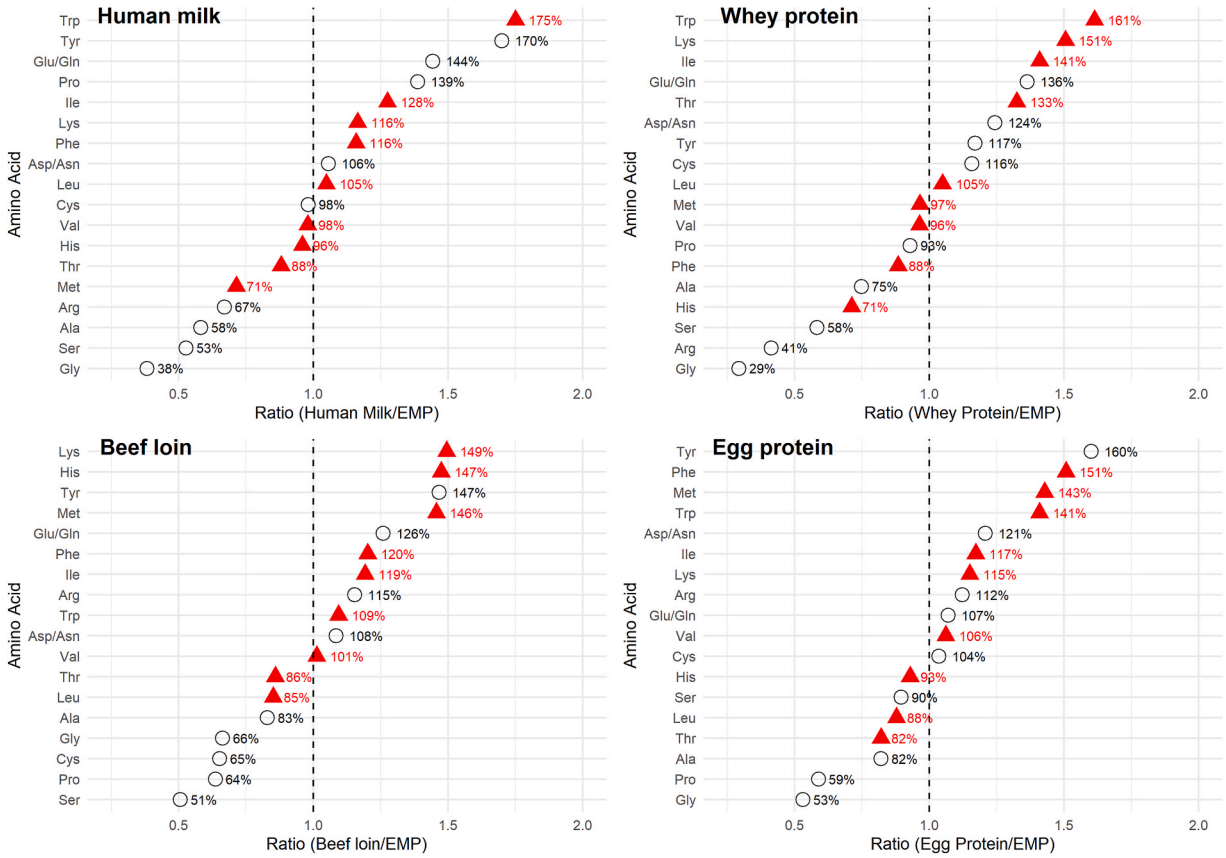


**Fig. 1.** A comparison between the concentration of amino acids (AAs) in animal foods and the human exome-matched protein (EMP) given in Table 2. The individual AAs are shown on the y-axis in descending order of the ratio between the concentration in the given protein and the concentration in EMP. The ratios are shown on the x-axis, with red triangles indicating the nine indispensable AAs and black circles the dispensable AAs. The horizontal line indicates a perfect match in the AA concentrations. The AA pattern for whey protein was taken from Morifuji et al. (2010), the pattern for human milk from Table 8 in Zhang et al. (2013), and the other patterns from Wu (2022). See Table 2 for the AA letter abbreviations. The plot has been produced in R version 4.5.2 with the ggplot2 and cowplot packages. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

out that the claim for 99% net nitrogen utilization would strictly be valid only for the very low intake of 0.4g MAP/kg/day as the sole protein source. Secondly, the small standard deviation of the nitrogen balance of 0.001 for the 66 subjects consuming MAP appears biologically and technically implausible, because this study applied the nitrogen balance method in which all nitrogen losses from sweat, feces, urine and other routes must be measured or otherwise accounted for, which is known to be a very difficult task and typically results in biased results (Wu, 2022). Thirdly, the fact that MAP is an incomplete protein lacking histidine and all 11 dispensable AAs would render it suboptimal as the sole protein source (Hou and Wu, 2017). While short-term studies by Lucà-Moretti et al. suggest that taking MAP as the sole protein source over several weeks is safe and beneficial in certain contexts (Lucà-Moretti et al., 2003a, 2003b), it would be expected that in the long term, diets composed of only indispensable AAs are suboptimal for growth, muscle mass acquisition and reproduction (Hou et al., 2016). Exactly these traits were shown to be optimized on an EMP diet in the preclinical studies discussed above (Gong et al., 2024; Ma et al., 2022; Piper et al., 2017; Wu et al., 2024), justifying further research to test whether the beneficial effects of EMPs translate to humans.

While MAP might still have value for supplying the body with highly bioavailable indispensable AAs,[11] EMP formulations appear more promising due to their completeness and deductively (error-free) derived target AA pattern. The ePROB algorithm developed by Monarch University (Bennett et al., 2023) principally allows the design of EMP powders, protein bars or enteral formulas from natural protein sources. One could envision several contexts under which individuals would potentially benefit from such supplements:

- Vegan diets: Plant-based proteins are generally inferior to animal-based proteins (Table 1), which is due to both their indispensable AA composition and poor digestibility because of antinutritional factors (Cordain, 1999; Gilani et al., 2012). Individuals on vegan diets may be at risk for protein deficiencies under certain contexts such as during growth (Dagnelie et al., 1989) or pregnancy. EMP supplements based on plant protein blends that have been processed by removing some of the antinutritional factors may be beneficial during a vegan diet, as they offer a simple way to improve both the quality and quantity of protein intake.
- Increased protein requirements: EMP supplements could benefit individuals whose protein needs are often difficult to meet through a standard diet, such as (elite) athletes, the elderly, or individuals suffering from chronic stress or illness.
- Chronic liver or kidney disease: Such individuals may benefit from reduced nitrogen waste production when replacing part of their standard dietary protein with EMP supplements and slightly reducing overall protein intake. This could allow for more efficient nitrogen utilization with less production of toxic ammonium as indicated by preclinical data (Piper et al., 2017).

[11] We have utilized a MAP supplement in addition to a ketogenic diet in a controlled clinical trial on breast cancer patients undergoing radiotherapy, but found no significant effect of taking MAP on body composition (Klement et al., 2020).

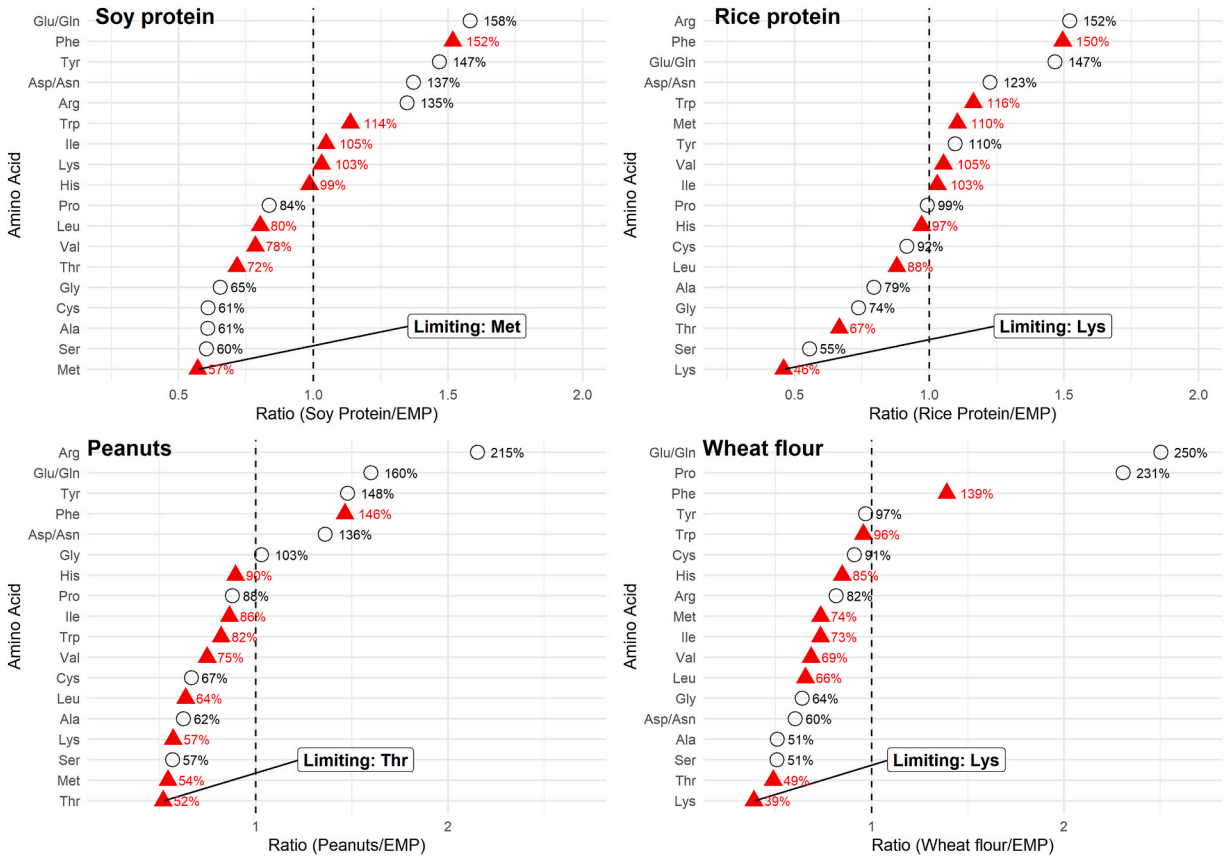


**Fig. 2.** Same as Fig. 1, but now for four different plant proteins. The AA pattern of soy protein was taken from Morifuji et al. (2010) and the other patterns from Wu (2022).

- Fasting Support: Alternatively, EMP supplements could be used during fasting periods to preserve valuable muscle mass while the body utilizes fats and ketone bodies as an energy source. Because protein ingestion in a glycogen-depleted state does not interfere with AMPK signaling (Taylor et al., 2013), carbohydrates are the major disturbers of fasting metabolism (Klein and Wolfe, 1992) and EMP is expected to yield particularly low amounts of AAs undergoing deamination and thus being precursors of gluconeogenesis, it can be hypothesized that the sole ingestion of an EMP supplement may not significantly interfere with the desired adaptions to fasting.

## 5. Conclusions and outlook

The concept of EMPs hypothetically defines an ideal reference protein composition, which can be applied to any species with available whole genome sequencing information. It is – as this review has tried to highlight – an exciting theoretical concept with the potential to stimulate further transdisciplinary research touching such disciplines as astrobiology, biochemistry, ecology and evolution, genetics, nutrition, physiology, veterinary medicine and economics. Prospects for further research regard the origin of and species-specific differences between exome-translated proteomes, their association with phylogenetic relationships among species and their relevance for species-specific protein quality metrics with practical applications in the livestock, pet feed and nutritional supplement industries. If translation of the preclinical findings turns out to be successful, the possibility emerges to improve the nutritional situation of large population cohorts with insufficient access to high-quality animal protein, or protein-rich foods in general, by the intelligent combination of different natural protein sources (Bennett et al., 2023). Overall, therefore, the concept of EMP diets appears promising and should be tested further in preclinical and clinical studies in the near future.

## Funding

This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Rainer Klement reports a relationship with MITOcare GmbH & Co. KG that includes: consulting or advisory and speaking and lecture fees.

## Data availability

No data was used for the research described in the article.

## References

Adhikari, S., Schop, M., de Boer, I.J.M., Huppertz, T., 2022. Protein quality in perspective: a review of protein quality metrics and their applications. Nutrients 14, 947. https://doi.org/10.3390/nu14050947.

Altwegg, K., Balsiger, H., Bar-Nun, A., Berthelier, J.J., Bieler, A., Bochsler, P., Briois, C., Calmonte, U., Combi, M.R., Cottin, H., De Keyser, J., Dhooghe, F., Fiethe, B., Fuselier, S.A., Gasc, S., Gombosi, T.I., Hansen, K.C., Haessig, M., Jäckel, A., Kopp, E., Korth, A., Le Roy, L., Mall, U., Marty, B., Mousis, O., Owen, T., Rème, H., Rubin, M., Sémon, T., Tzou, C.Y., Waite, J.H., Wurz, P., 2016. Prebiotic chemicals-amino acid and phosphorus in the coma of comet 67P/Churyumov-Gerasimenko. Sci. Adv. 2, e1600285. https://doi.org/10.1126/sciadv.1600285.

Appler, K.E., Lingford, J.P., Gong, X., Panagiotou, K., Leão, P., Langwig, M.V., Greening, C., Ettema, T.J.G., Anda, V. De, Baker, B.J., 2026. Oxygen metabolism in descendants of the archaeal-eukaryotic ancestor. Nature 652, 405–415. https://doi.org/10.1038/s41586-026-10128-z.

Arends, J., Bachmann, P., Baracos, V., Barthelemy, N., Bertz, H., Bozzetti, F., Fearon, K., Hütterer, E., Isenring, E., Kaasa, S., Krznaric, Z., Laird, B., Larsson, M., Laviano, A., Mühlebach, S., Muscaritoli, M., Oldervoll, L., Ravasco, P., Solheim, T., Strasser, F., de van der Schueren, M., Preiser, J.C., 2017. ESPEN guidelines on nutrition in cancer patients. Clin. Nutr. 36, 11–48. https://doi.org/10.1016/j.clnu.2016.07.015.

Aspden, J.L., Wallace, E.W.J., Whiffin, N., 2023. Not all exons are protein coding: addressing a common misconception. Cell Genom. 3, 100296. https://doi.org/10.1016/j.xgen.2023.100296.

Bauer, J., Biolo, G., Cederholm, T., Cesari, M., Cruz-Jentoft, A.J., Morley, J.E., Phillips, S., Sieber, C., Stehle, P., Teta, D., Visvanathan, R., Volpi, E., Boirie, Y., 2013. Evidence-based recommendations for optimal dietary protein intake in older people: a position paper from the prot-age study group. J. Am. Med. Dir. Assoc. 14, 542–559. https://doi.org/10.1016/j.jamda.2013.05.021.

Bennett, L., Piper, M., Ernst, A., Wu, T., 2023. ePROB: a new tool for defining protein quality. Food Aust. 75, 19–21. https://doi.org/10.3316/informit.858843876174869.

Burd, N.A., McKenna, C.F., Salvador, A.F., Paulussen, K.J.M., Moore, D.R., 2019. Dietary protein quantity, quality, and exercise are key to healthy living: a muscle-centric perspective across the lifespan. Front. Nutr. 6, 83. https://doi.org/10.3389/fnut.2019.00083.

Commisso, C., Davidson, S.M., Soydaner-Azeloglu, R.G., Parker, S.J., Kamphorst, J.J., Hackett, S., Grabocka, E., Nofal, M., Drebin, J.A., Thompson, C.B., Rabinowitz, J.D., Metallo, C.M., Vander Heiden, M.G., Bar-Sagi, D., 2013. Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. Nature 497, 633–637. https://doi.org/10.1038/nature12138.

Cordain, L., 1999. Cereal grains: humanity's double-edged sword. In: Simopoulos, A.P. (Ed.), World Review of Nutrition and Dietetics. Karger, Basel, pp. 19–73.

Dagnelie, P.C., van Staveren, W.A., Verschuren, S.A., Hautvast, J.G., 1989. Nutritional status of infants aged 4 to 18 months on macrobiotic diets and matched omnivorous control infants: a population-based mixed-longitudinal study. I. Weaning pattern, energy and nutrient intake. Eur. J. Clin. Nutr. 43, 311–323.

Duan, Y., Cao, Q., 2025. Systematic revelation and meditation on the significance of long exons using representative eukaryotic genomes. BMC Genom. 26, 290. https://doi.org/10.1186/s12864-025-11504-1.

Elango, R., Ball, R.O., Pencharz, P.B., 2008. Indicator amino acid oxidation: concept and application. J. Nutr. 138, 243–246. https://doi.org/10.1093/jn/138.2.243.

Elsila, J.E., Glavin, D.P., Dworkin, J.P., 2009. Cometary glycine detected in samples returned by Stardust. Meteorit. Planet. Sci. 44, 1323–1330. https://doi.org/10.1111/j.1945-5100.2009.tb01224.x.

Esumi, G., 2025. Chicken eggs are a practical and common exome-matched diet for multicellular eukaryotic organisms. Jxiv 1056. https://doi.org/10.51094/jxiv.1056.

FAO, 2013. Dietary protein quality evaluation in human nutrition. Report of an FAQ expert consultation. FAO Food Nutr. Pap.

Gilani, G.S., Xiao, C.W., Cockell, K.A., 2012. Impact of antinutritional factors in food proteins on the digestibility of protein and the bioavailability of amino acids and on protein quality. Br. J. Nutr. 108, S315–S332. https://doi.org/10.1017/S0007114512002371.

Glavin, D.P., Dworkin, J.P., Aubrey, A., Botta, O., Doty, J.H., Martins, Z., Bada, J.L., 2006. Amino acid analyses of antarctic CM2 meteorites using liquid chromatography-time of flight-mass spectrometry. Meteorit. Planet. Sci. 41, 889–902. https://doi.org/10.1111/j.1945-5100.2006.tb00493.x.

Goldman, A.D., Becerra, A., 2024. A new view of the last universal common ancestor. J. Mol. Evol. 92, 659–661. https://doi.org/10.1007/s00239-024-10193-w.

Gong, C., Jiao, C., Liang, H., Ma, Y., Wu, Q., Xie, Y., 2024. Exome-based amino acid optimization: a dietary strategy to satisfy human nutritional demands and enhance muscle strength in breast tumor mice undergoing chemotherapy. J. Agric. Food Chem. 72, 7089–7099. https://doi.org/10.1021/acs.jafc.3c07256.

Halliday, A.N., Canup, R.M., 2023. The accretion of planet Earth. Nat Rev Earth Environ 4, 19–35. https://doi.org/10.1038/s43017-022-00370-0.

Herreman, L., Nommensen, P., Pennings, B., Laus, M.C., 2020. Comprehensive overview of the quality of plant- and animal-sourced proteins based on the digestible indispensable amino acid score. Food Sci. Nutr. 8, 5379–5391. https://doi.org/10.1002/fsn3.1809.

Hirshowitz, B., Brook, J., Kaufman, T., Titelman, U., Mahler, D., 1975. 35 eggs per day in the treatment of severe burns. Br. J. Plast. Surg. 28, 185–188.

Hou, Y., Wu, G., 2017. Nutritionally nonessential amino acids: a misnomer in nutritional sciences. Adv. Nutr. 8, 137–139. https://doi.org/10.3945/an.116.012971.

Hou, Y., Yao, K., Yin, Y., Wu, G., 2016. Endogenous synthesis of amino acids limits growth, lactation, and reproduction in animals. Adv. Nutr. 7, 331–342. https://doi.org/10.3945/an.115.010850.

Hou, Y., Yin, Y., Wu, G., 2015. Dietary essentiality of "nutritionally non-essential amino acids" for animals and humans. Exp Biol Med 240, 997–1007. https://doi.org/10.1177/1535370215587913.

Iannotti, L.L., Lutter, C.K., Stewart, C.P., Riofrío, C.A.G., Malo, C., Reinhart, G., Palacios, A., Karp, C., Chapnick, M., Cox, K., Waters, W.F., 2017. Eggs in early complementary feeding and child growth: a randomized controlled trial. Pediatrics 140, e20163459. https://doi.org/10.1542/peds.2016-3459.

Iglesias-Groth, S., Dobrincic, M.M., Garrido, A.P., 2025. Evidence of the amino acids tyrosine and phenylalanine in the interstellar material of IC348 in Perseus. Life 15, 181. https://doi.org/10.3390/life15020181.

Ioppolo, S., Fedoseev, G., Chuang, K.J., Cuppen, H.M., Clements, A.R., Jin, M., Garrod, R. T., Qasim, D., Kofman, V., van Dishoeck, E.F., Linnartz, H., 2021. A non-energetic mechanism for glycine formation in the interstellar medium. Nat. Astron. 5, 197–205. https://doi.org/10.1038/s41550-020-01249-0.

Kirschning, A., 2022. On the evolutionary history of the twenty encoded amino acids. Chem. Eur J. 28, e202201419. https://doi.org/10.1002/chem.202201419.

Klein, S., Wolfe, R.R., 1992. Carbohydrate restriction regulates the adaptive response to fasting. Am. J. Physiol. 262, E631–E636. https://doi.org/10.1152/ajpendo.1992.262.5.E631.

Klement, R.J., 2024. The optimal amino acid pattern for humans and its implications for nutrition of cancer patients. Transl Breast Cancer Res 5, 25. https://doi.org/10.21037/tbcr-24-25.

Klement, R.J., Champ, C.E., Kämmerer, U., Koebrunner, P.S., Krage, K., Schäfer, G., Weigel, M., Sweeney, R.A., 2020. Impact of a ketogenic diet intervention during radiotherapy on body composition: III—final results of the KETOCOMP study for breast cancer patients. Breast Cancer Res. 22, 94. https://doi.org/10.1186/s13058-020-01331-5.

Koonin, E.V., 2009. Intron-dominated genomes of early ancestors of eukaryotes. J. Hered. 100, 618–623. https://doi.org/10.1093/jhered/esp056.

Koonin, E.V., Novozhilov, A.S., 2017. Origin and evolution of the universal genetic code. Annu. Rev. Genet. 51, 45–62. https://doi.org/10.1146/annurev-genet-120116-024713.

Krasnokutski, S.A., Chuang, K.J., Jäger, C., Ueberschaar, N., Henning, T., 2022. A pathway to peptides in space through the condensation of atomic carbon. Nat. Astron. 6, 381–386. https://doi.org/10.1038/s41550-021-01577-9.

Lucà-Moretti, M., 1998. A comparative, double-blind, triple crossover net nitrogen utilization study confirms the discovery of the master amino acid pattern. Ann R Natl Acad Med Spain CXV, 397–416.

Lucà-Moretti, M., Grandi, A., Lucà, E., Mariani, E., Vender, G., Arrigiotti, E., Ferrario, M., Rovelli, E., 2003a. Results of taking master amino acid pattern as a sole and total substitute of dietary proteins in an athlete during a desert crossing. Adv. Ther. 20, 203–210.

Lucà-Moretti, M., Grandi, A., Lucà, E., Muratori, G., Nofroni, M., Mucci, M., Gambetta, P., Stimolo, R., Drago, P., Giudice, G., Tamburlin, N., Karbalai, M., Valente, C., Moras, G., 2003b. Master amino acid pattern as sole and total substitute for dietary proteins during a weight-loss diet to achieve the body's nitrogen balance equilibrium. Adv. Ther. 20, 270–281.

Ma, C., Mirth, C.K., Hall, M.D., Piper, M.D.W., 2022. Amino acid quality modifies the quantitative availability of protein for reproduction in Drosophila melanogaster. J. Insect Physiol. 139, 104050. https://doi.org/10.1016/j.jinsphys.2020.104050.

MacArthur, M.R., Mitchell, J.R., 2017. Feeding the genome: in silico optimization of dietary amino acid composition. Cell Metab. 25, 486–488. https://doi.org/10.1016/j.cmet.2017.02.015.

Makarov, M., Rocha, A.C.S., Krystufek, R., Cherepashuk, I., Dzmitruk, V., Charnavets, T., Faustino, A.M., Lebl, M., Fujishima, K., Fried, S.D., Hlouchova, K., 2023. Early selection of the amino acid alphabet was adaptively shaped by biophysical constraints of foldability. J. Am. Chem. Soc. 145, 5320–5329. https://doi.org/10.1021/jacs.2c12987.

Martin, W.F., 2020. Older than genes: the acetyl CoA pathway and origins. Front. Microbiol. 11, 817. https://doi.org/10.3389/fmicb.2020.00817.

Mayer-Bacon, C., Freeland, S.J., 2021. A broader context for understanding amino acid alphabet optimality. J. Theor. Biol. 520, 110661. https://doi.org/10.1016/j.jtbi.2021.110661.

Moody, E.R.R., Álvarez-Carretero, S., Mahendrarajah, T.A., Clark, J.W., Betts, H.C., Dombrowski, N., Szánthó, L.L., Boyle, R.A., Daines, S., Chen, X., Lane, N., Yang, Z., Shields, G.A., Szöllősi, G.J., Spang, A., Pisani, D., Williams, T.A., Lenton, T.M., Donoghue, P.C.J., 2024. The nature of the last universal common ancestor and its impact on the early Earth system. Nat. Ecol. Evol. 8, 1654–1666. https://doi.org/10.1038/s41559-024-02461-1.

Morifuji, M., Ishizaka, M., Baba, S., Fukuda, K., Matsumoto, H., Koga, J., Kanegae, M., Higuchi, M., 2010. Comparison of different sources and degrees of hydrolysis of dietary protein: effect on plasma amino acids, dipeptides, and insulin responses in human subjects. J. Agric. Food Chem. 58, 8788–8797. https://doi.org/10.1021/jf101912n.

Munñoz Caro, G.M., Melerhenrich, U.J., Schutte, W.A., Barbier, B., Arcones Segovia, A., Rosenbauer, H., Thiemann, W.H.P., Brack, A., Greenberg, J.M., 2002. Amino acids from ultraviolet irradiation of interstellar ice analogues. Nature 416, 403–406. https://doi.org/10.1038/416403a.

Ortega, J.G., Raubenheimer, D., Tyagi, S., Mirth, C.K., Piper, M.D.W., 2023. Biosynthetic constraints on amino acid synthesis at the base of the food chain may determine their use in higher-order consumer genomes. PLoS Genet. 19, e1010635. https://doi.org/10.1371/journal.pgen.1010635.

Philip, G.K., Freeland, S.J., 2011. Did evolution select a nonrandom "alphabet" of amino acids? Astrobiology 11, 235–240. https://doi.org/10.1089/ast.2010.0567.

Piper, M.D.W., Soultoukis, G.A., Blanc, E., Mesaros, A., Herbert, S.L., Juricic, P., He, X., Atanassov, I., Salmonowicz, H., Yang, M., Simpson, S.J., Ribeiro, C., Partridge, L., 2017. Matching dietary amino acid balance to the in silico-translated exome optimizes growth and reproduction without cost to lifespan. Cell Metab. 25, 610–621. https://doi.org/10.1016/j.cmet.2017.02.005.

Poverennaya, I.V., Roytberg, M.A., 2020. Spliceosomal introns: features, functions, and evolution. Biochem 85, 725–734. https://doi.org/10.1134/S0006297920070019.

Rogozin, I.B., Carmel, L., Csuros, M., Koonin, E.V., 2012. Origin and evolution of spliceosomal introns. Biol. Direct 7, 1–28. https://doi.org/10.1186/1745-6150-7-11.

Rotz, C.A., 2004. Management to reduce nitrogen losses in animal production. J. Anim. Sci. 82, E119–E137. https://doi.org/10.2527/2004.8213_supplE119x.

Ruiz-Lara, T., Gallart, C., Bernard, E.J., Cassisi, S., 2020. The recurrent impact of the Sagittarius dwarf on the star formation history of the Milky Way. Nat. Astron. 4, 965–973. https://doi.org/10.1038/s41550-020-1097-0.

Sakharkar, M.K., Chow, V.T.K., Kangueane, P., 2004. Distributions of exons and introns in the human genome. In Silico Biol. 4, 387–393. https://doi.org/10.3233/isb-00142.

Seyfried, T.N., Chinopoulos, C., 2021. Can the mitochondrial metabolic theory explain better the origin and management of cancer than can the somatic mutation theory? Metabolites 11, 572. https://doi.org/10.3390/metabo11090572.

Shimoyma, A., Ogasawara, R., 2002. Dipeptides and Diketopiperazines in the Yamato-791198 and Murchison carbonaceous chondrites. Orig. Life Evol. Biosph. 32, 165–179. https://doi.org/10.1023/A:1016015319112.

Shivani, P.P., Misra, A., Tandon, P., 2017. A theoretical quantum chemical study of alanine formation. Eur. Phys. J. D 71, 215. https://doi.org/10.1140/epjd/e2017-70575-2.

Steele, E.J., Gorczynski, R.M., Lindley, R.A., Liu, Y., Temple, R., Tokoro, G., Wickramasinghe, D.T., Wickramasinghe, N.C., 2019. Lamarck and panspermia - on the efficient spread of living systems throughout the cosmos. Prog. Biophys. Mol. Biol. 149, 10–32. https://doi.org/10.1016/j.pbiomolbio.2019.08.010.

Stepankiw, N., Yang, A.W.H., Hughes, T.R., 2023. The human genome contains over a million autonomous exons. Genome Res. 33, 1865–1878. https://doi.org/10.1101/gr.277792.123.

Taylor, C., Bartlett, J.D., van de Graaf, C.S., Louhelainen, J., Coyne, V., Iqbal, Z., Maclaren, D.P.M., Gregson, W., Close, G.L., Morton, J.P., 2013. Protein ingestion does not impair exercise-induced AMPK signalling when in a glycogen-depleted state: implications for train-low compete-high. Eur. J. Appl. Physiol. 113, 1457–1468. https://doi.org/10.1007/s00421-012-2574-7.

Wehbi, S., Wheelr, A., Morel, B., Manepalli, N., Minh, B.Q., Lauretta, D.S., Masel, J., 2024. Order of amino acid recruitment into the genetic code resolved by last universal common ancestor's protein domains. Proc. Natl. Acad. Sci. USA 121, e2410311121. https://doi.org/10.1073/pnas.24103111211.

Wickramasinghe, C., 2018. Proofs that Life is Cosmic, first ed. World Scientific Publishing Co Pte Ltd, Singapore.

Woon, D.E., 2002. Pathways to glycine and other amino acids in ultraviolet-irradiated astrophysical ices determined via Quantum chemical modeling. Astrophys. J. 571, L177–L180. https://doi.org/10.1086/341227.

World Health Organization, 2007. Protein and Amino Acid Requirements in Human Nutrition: Report of a Joint FAO/WHO/UNU Expert Consultation, WHO Technical Report Series. WHO Press, Geneva, Switzerland.

Wu, G., 2022. Amino Acids. Biochemistry and Nutrition, second ed. CRC Press, Boca Raton, FL. https://doi.org/10.1201/9781003092742.

Wu, T., Baatar, D., O' Connor, A.E., O'Bryan, M.K., Stringer, J.M., Hutt, K.J., Malimige Aponso, M., Monro, K., Luo, J., Zhu, Y., Ernst, A., Swindells, E.O.K., Alesi, L.R., Tho Tony Nguyen, N., Piper, M.D.W., Bennett, L.E., 2024. Exome-informed formulations of food proteins enhance body growth and feed conversion efficiency in ad libitum-fed mice. Food Res. Int. 176, 113819. https://doi.org/10.1016/j.foodres.2023.113819.

Wu, T., Taylor, C., Nebl, T., Ng, K., Bennett, L.E., 2017. Effects of chemical composition and baking on in vitro digestibility of proteins in breads made from selected gluten-containing and gluten-free flours. Food Chem. 233, 514–524. https://doi.org/10.1016/j.foodchem.2017.04.158.

Zhang, C., Wang, C., Li, Y., Miwa, T., Liu, C., Cui, W., Song, W.C., Du, J., 2017. Complement C3a signaling facilitates skeletal muscle regeneration by regulating monocyte function and trafficking. Nat. Commun. 8, 2078. https://doi.org/10.1038/s41467-017-01526-z.

Zhang, Z., Adelman, A.S., Rai, D., Boettcher, J., Lonnerdal, B., 2013. Amino acid profiles in term and preterm human milk through lactation: a systematic review. Nutrients 5, 4800–4821. https://doi.org/10.3390/nu5124800.

Zimmerly, S., Semper, C., 2015. Evolution of group II introns. Mob DNA 6, 7. https://doi.org/10.1186/s13100-015-0037-5.